U.fm

Aprendizagem Computacional *

Início	quarta, 9 de fevereiro de 2022 às 19:50
Estado	Prova submetida
Data de submissão:	quarta, 9 de fevereiro de 2022 às 19:50
Tempo gasto	7 segundos
Nota	0,0/12,0
Nota	**0,0** de um máximo de 20,0 (**0**%)

Pergunta 1

Não respondida Pontuação 1,0 Destacar pergunta

From the broadest possible perspective, data mining approaches analyze data obtained from a phenomenon to:

a. understand or anticipate future deviations from expected behavior.
b. understand past behavior or predict future behavior.
c. Não quero responder
d. predict future behavior and explain the causes of that behavior.
e. anticipate future deviations from expected behavior and explain the causes of these deviations.

A resposta correta é: understand past behavior or predict future behavior.

Pergunta 2

Não respondida Pontuação 1,0 Destacar pergunta

In general, the most important challenge in carrying out a successful data mining project is the lack of guarantees regarding the results because:

a. The data may not contain the information needed to achieve the business objectives.
b. Não quero responder
c. CRISP-DM methodology does not provide instruments to validate these results.
d. the volume of data is so large that it is not possible to analyze it.
e. The CRISP-DM methodology does not guarantee that the business objectives are quantifiable.

A resposta correta é: The data may not contain the information needed to achieve the business objectives.

Pergunta 3

Não respondida Pontuação 1,0 Destacar pergunta

Indicate in which of the following cases the reliability of the results of a poll on the vote intentions would be significantly affected. The poll is made on the basis of a sample of citizens:

a. Não quero responder
b. national and/or foreigners entitled to vote.
c. selected completely randomly.
d. with minimum age to vote, in accordance with the electoral law of Portugal.
e. with landline.

A resposta correta é: with minimum age to vote, in accordance with the electoral law of Portugal.

Pergunta 4

Não respondida Pontuação 1,0 Destacar pergunta

A set of data has missing values in a numeric variable, caused by data collection errors. Which of the following techniques is best suited to fill these values:
Selecione uma opção de resposta:

a. the most frequent observed value.
b. median.
c. Não quero responder
d. a value much higher than the highest observed value.
e. 0.

A resposta correta é: median.

Pergunta 5

Não respondida Pontuação 1,0 Destacar pergunta

In a classification problem, where the objective is to predict whether a bus will be delayed in arriving at the next stop or not, the use of the data to estimate the generalization error should be:

a. use a random sample of 70 percent of the data for training and the rest for testing.
b. given day d, use data from d+1 for training and data up to d for testing.
c. Não quero responder
d. use a random sample of 30 percent of the data for training and the rest for testing.
e. given day d, use data up to d-1 for training and data from d for testing.

A resposta correta é: given day d, use data up to d-1 for training and data from d for testing.

Pergunta 6

Não respondida Pontuação 1,0 Destacar pergunta

A prediction problem can be seen as, given a phenomenon y=f(x1, x2, ..., xN), apply a learning algorithm:

a. to evaluate whether the value of the function f(x1, x2, ..., xN) is really y.
b. to the training data, represented as (y, x1, x2, ..., xN), to try to identify f.
c. to the training data, represented as (y, x1, x2, ..., xN), to assess whether f is really the function that determines the relationship between the target variable y and the independent variables (x1, x2 , ..., xN).
d. to the training data, represented as f(y, x1, x2, ..., xN), to try to predict the value of y.
e. Não quero responder

A resposta correta é: to the training data, represented as (y, x1, x2, ..., xN), to try to identify f.

Pergunta 7

Não respondida Pontuação 1,0 Destacar pergunta

In the k-means algorithm, as the value of the input parameter k increases, the sum of the within clusters quadratic errors:

a. decrease.
b. does not change.
c. may increase or decrease.
d. increase.
e. Não quero responder

A resposta correta é: decrease.

Pergunta 8

Não respondida Pontuação 1,0 Destacar pergunta

If the rule {potatos}=>{carrots} has a support of 0.5 and confidence of 1.0 then the rule {carrots}=>{lettuce} will have a support of:

a. this case is impossible.
b. 0.5 or higher.
c. 0.5.
d. less than 0.5.
e. Não quero responder

A resposta correta é: 0.5 or higher.

Pergunta 9

Não respondida Pontuação 1,0 Destacar pergunta

One method to control overfitting in decision trees is to:

a. reduce the number of levels in the tree, to minimize the probability that a leaf will have an insufficient number of examples.
b. increase the number of levels in the tree, to minimize the probability that a leaf will have an excessive number of examples.
c. increase the number of levels in the tree, to minimize the probability that a leaf will have an insufficient number of examples.
d. Não quero responder
e. reduce the number of levels in the tree, to minimize the probability that a leaf will have an excessive number of examples.

A resposta correta é: reduce the number of levels in the tree, to minimize the probability that a leaf will have an insufficient number of examples.

Pergunta 10

Não respondida Pontuação 1,0 Destacar pergunta

The AUC measure:

a. Não quero responder
b. represents graphically the probability that a negative example is ranked higher than a positive example as the decision threshold decreases.
c. quantifiesthe probability that a positive example is ranked higher than a negative example.
d. quantifies the probability that a negative example is ranked higher than a positive example.
e. represents graphically the probability that a positive example is ranked higher than a negative example as the decision threshold decreases.

A resposta correta é: quantifies the probability that a negative example is ranked higher than a positive example.

Pergunta 11

Não respondida Pontuação 1,0 Destacar pergunta

A Neural Network is a universal approximator because:

a. it can model any continuous function with only 1 node in the internal layer.
b. it can model any continuous function, given a sufficient number of nodes.
c. Não quero responder
d. it can learn any model that any other algorithm learns, including decision trees, SVM and naive Bayes.
e. it can learn any model that any other algorithm learns, including decision trees, SVM and naive Bayes, given a sufficient number of nodes.

A resposta correta é: it can model any continuous function, given a sufficient number of nodes.

Pergunta 12

Não respondida Pontuação 1,0 Destacar pergunta

Metalearning can be described as:

a. a learning approach to understand the behavior of learning algorithms.
b. the development of super-algorithms that are immune to the curse of dimensionality.
c. Não quero responder
d. the development of super-algorithms that are immune to the No-Free-Lunch theorem.
e. a learning approach to correct the predictions made by learning algorithms.

A resposta correta é: a learning approach to understand the behavior of learning algorithms.